DOCUMENT RESUME

ED 151 399 TM 006 861

| | |
|---|---|
| TITLE | Evaluation Design: New York State Experimental Prekindergarten Program. |
| INSTITUTION | New York State Education Dept., Albany. Bureau of Child Development and Parent Education. |
| PUB DATE | 3 Sep 75 |
| NOTE | 32p.; For related documents, see TM 006 634, 862, 863, and 903 |
| EDRS PRICE | MF-$0.83 HC-$2.06 Plus Postage. |
| DESCRIPTORS | Affective Tests; *Cognitive Development; Cognitive Tests; Compensatory Education Programs; Data Analysis; Data Collection; Disadvantaged Youth; Early Childhood Education; *Early Experience; Family Characteristics; Individual Development; *Longitudinal Studies; Parental Background; Parent Participation; Preschool Evaluation; *Preschool Programs; Preschool Tests; *Program Effectiveness; Program Length; *Research Design; Socioeconomic Status; Student Characteristics |
| IDENTIFIERS | California Preschool Social Competency Scale; Classroom Behavior Inventory (Schaefer et al); Cooperative Preschool Inventory; New York; New York Experimental Prekindergarten Program; Peabody Picture Vocabulary Test; Walker Readiness Test |

ABSTRACT

In order to expose disadvantaged preschool children to a variety of educational experiences and to health and social services, the New York State Legislature funded the State Experimental Prekindergarten Program (PreK). In 1975, a five-year longitudinal evaluation study was begun. The study has two major parts: (1) a general study of 5,800 students in 53 prekindergarten programs, continuing through the third grade, and (2) an in-depth study of 800 students in the early childhood programs of seven districts, also continuing through grade 3. The general study investigates the impact of program variations, particularly the number of contact hours, home vs. center based programs, extent of parent involvement, and teacher variables, on students with different cognitive and affective characteristics. The in-depth study documents, through systematic observation and record keeping, the continuity and characteristics of the preschool experiences. (MV)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED151399

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

David J. Irvine

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

# EVALUATION DESIGN

New York State

Experimental Prekindergarten Program

September 3, 1975

The University of the State of New York
THE STATE EDUCATION DEPARTMENT
Bureau of Child Development and Parent Education
Bureau of School Programs Evaluation
Albany, New York 12234

TM006 861

## INTRODUCTION

Experiences in the early years play an important role in determining a child's future development. Some children in our society do not receive the variety or types of experiences which are a part of the backgrounds of many children, with the result that these children often have difficulty coping with requirements placed on them in school and in other aspects of their lives. Exposing them to a variety of educational experiences and comprehensive health and social services before they start the traditional school program could help reduce or eliminate this difficulty. To provide such experiences and services, the New York State Legislature initiated annual funding of State Experimental Prekindergarten (Pre-K) Program under Chapter 46 of the Laws of 1966.

## OVERVIEW

Several questions to be answered by this project are listed below:

1. Do Prekindergarten programs have a measurable effect upon children's cognitive and noncognitive development between the time they enter and leave the program?

This question will be answered by administering cognitive measures [Peabody Picture Vocabulary Test (PPVT), Walker Readiness Test (WRT), and the Cooperative Preschool Inventory (CPI)] and affective measures [Classroom Behavior Inventory (CBI) and California Preschool Social Competency Scale (CPSCS)] both at the beginning and end of the prekindergarten year. Two types of comparisons will be made. First, a small control group of approximately 100 children will be selected from the

district's lists of children who were eligible for inclusion in the program but for various reasons could not be accomodated. Since the two affective measures are teacher ratings of pupil behavior which require the child to be in a group setting, the control group will be tested on only the three cognitive measures. Testing of the control group will take place within the same time frame as that of the group attending prekindergarten. Direct comparisons between treatments (prekindergarten vs. no prekindergarten) will be made.

Second, scores on two of the cognitive measures (Walker Readiness and Cooperative Preschool Inventory) will be compared to their respective norm group with some confidence since both tests were normed on children from low socioeconomic families. However, differences between pretest and posttest scores on the third cognitive measure (PPVT) and on the affective measures will be difficult to interpret using the tests' norm groups. These differences between pre- and posttesting might be attributed either to the prekindergarten program or to general development.

In both cases, an analysis of covariance using the Walker Test as a covariate will be employed to analyze the data.

2. Do prekindergarten children possess measurable advantages in cognitive development over non-prekindergarten children at the time of entry to kindergarten?

This question will be answered by comparing children with prekindergarten experience to children without prekindergarten experience at the beginning of the kindergarten year on the same measures mentioned above. This control group will be selected from districts in which there was no

Experimental Prekindergarten Program operating and which are similar to the districts which sponsor prekindergarten programs. Students in this control group will be sampled from a population which is comparable to the prekindergarten population in socioeconomic background. Bronfenbrenner (1974) and others have suggested that socioeconomic status (SES) variables seem to contribute quite heavily to the longitudinal effects of prekindergarten programs. Therefore, comparable groups (based on SES) would be essential in obtaining a valid evaluation. However, there is some concern whether obtaining a comparable group will be possible since in addition to the Experimental Prekindergarten Program there is a myriad of publicly funded programs for low socioeconomic children operating in the State. It may be difficult to locate a sufficient number of children with similar socioeconomic backgrounds who have not participated in a preschool program. In view of the anticipated difficulty, the small control group identified during the prekindergarten year will be of great assistance when selecting the large control group at the beginning of the kindergarten year with respect to the appropriate criteria for selection.

3. Does prekindergarten experience have a measurable effect of children's cognitive development through the time they complete the third grade?

The question will be answered by testing the same groups of children (with and without prekindergarten experience) at the end of each grade through the third grade. The control group will be the same as described under question 2. Significant differences between the two groups will be tested by analysis of covariance, multivariate analysis of variance, and

multiple regression analysis.

Two of the measures, the Walker Readiness Test and the Cooperative Preschool Inventory, are not normed beyond age 7. At the beginning of second grade, an achievement test series such as the Metropolitan Achievement Tests (MAT) will replace these cognitive measures. The MAT correlates very highly with the Walker and appears to be a reasonable replacement.

4. Are there specific variations in prekindergarten programs which result in significantly different levels of cognitive and other development in their students?

Program variation -- amount of exposure time (half day or whole day by the number of days per week),whether the program was center or home based, and amount of parental involvement -- will serve as treatment variables in the analysis. These treatment variables as well as the SES measure and initial level of performance on the Walker will be related to the three cognitive measures at the end of the prekindergarten program . This procedure will make it possible to determine the effect of different program variations on the posttest scores of students and whether these program variations interact with the SES variable and different initial levels of performance as measured by the pretest scores.

5. Can program components, or variations be identified which contribute to greater success of particular programs in producing lasting increases in the level of cognitive and other development?

Answers to this question will be obtained following the procedures described under question 4, with the following exceptions: a) The time

period will extend over prekindergarten, kindergarten, and the primary grades, with posttesting at the beginning and end of prekindergarten and kindergarten and at the end of each grade through third grade; and b) Program variations will also include different experiences students have in kindergarten and grades one through three. Special attention will be given to articulation (or lack of it) between the Prekindergarten program and the subsequent regular school program.

6. Will the systematic collection and analysis of data, carried out in examining the previous questions, provide a baseline for decision makers in studying staff, schedules, objectives, materials, child and parent characteristics, and children's mental ability and cognitive development?

As data are collected to answer the first five questions, attention will also be given to how the mechanisms for collecting data and the data themselves can serve the information needs of decision makers aside from the purposes of the Experimental Prekindergarten Program. The purpose will be to develop data and data-collection procedures which make it possible to provide useful and reliable data for making decisions about educational programs.

## RESEARCH DESIGN

The study will have two major parts:

1. A general study of 53 prekindergarten programs which will continue until the pupils complete grade 3 (approximately 5800 pupils).

2. An in-depth study of the prekindergarten and early elementary

programs of seven districts which will continue until the pupils complete grade 3 (approximately 800 pupils). The in-depth study will utilize the special skills of the Field Coordinator and attempt to document through systematic observation and record keeping the educational processes to which pupils are exposed.

The General Study

All of the 4-year-old students in the 53 prekindergarten programs will be tested at the start of the program. Additional information on students such as SES variables will be collected (see Initial Pupil Information form attached) as well as parental involvement as the project progresses.

A sample of approximately 1300 students will be selected during the prekindergarten year and tested at each of several critical points over the succeeding five years: at the end of prekindergarten, at the beginning and end of kindergarten, and at the end of each school year through grade 3. The sample will be representative of the original population with respect to initial cognitive scores, affective scores, and socioeconomic status and will be drawn using a sampling plan developed by the Bureau of Statistical Services. Children who are not available for testing at each of these points will be randomly replaced for cross-sectional analysis. The sample will make it possible to generalize to the population within 5 percentage points with 95 percent confidence.

A small control group will be selected and tested during the prekindergarten year in order to establish the effect of the program during the first year. Program directors will be asked to identify children

for the control group who were eligible for the prekindergarten program but for various reasons could not participate. State Education Department staff will be responsible for the data collection on as many of those children as possible.

For the control group, data will be collected only on the cognitive domain since noncognitive data consists of teacher ratings. Pretest data for the control group will be compared to pretest data for the experimental prekindergarten group. If the distribution of pretest scores for the two groups are significantly different, adjustments will be made on the control group to make it as similar to the prekindergarten group as possible. This will substantially reduce the likelihood of making inaccurate statements about the prekindergarten effect. Since the cognitive measures are developmental in nature, the data collected on these children will be available to the kindergarten teachers for diagnostic purposes the following year. In addition, characteristics of this small control group will be used as a basis for determining the appropriate criteria for the selection of the large control group at the beginning of kindergarten.

If no control group were available until kindergarten, certain assumptions would be implicit in the criteria for selection of the larger control group. Should SES variables be the selection criteria, an implicit assumption is that children of the same SES level function in an equivalent manner on cognitive tasks. This is an empirical question

and is not an assumption upon which a research design should be based. If equivalent cognitive skills at the kindergarten level are the selection criteria, children will have been matched on the criteria that should reflect differences. Kindergarteners who took part in a prekindergarten program would not be expected to behave either cognitively or affectively like the control group. If they did behave in a like manner, it would suggest that the prekindergarten programs were not effective.

To make it possible to determine whether gains made after prekindergarten might have occured without the prekindergarten experience, the second control group of 1300 beginning kindergarten students will be drawn from districts where no prekindergarten programs are operating, again according to a plan developed by the Bureau of Statistical Services. Students from those districts will be selected according to the criteria used to select students for prekindergarten programs so that the non-prekindergarten students match the prekindergarten students. It is not possible to draw a comparable control group from these districts participating in the experimental prekindergarten program because the neediest children meeting the criteria for admission were selected for the program. Non-prekindergarten students in the prekindergarten districts would not provide an equivalent group. While this control group does not afford comparison of the two groups during prekindergarten, it will provide a means of assessing the progress of prekindergarten students in kindergarten and grades one through three.

The focus of the evaluation in the general study will be to determine the impact of program variations, especially in the areas of number of contact hours, home vs. center based, extent of parental involvement, and teacher differences on students with different cognitive and affective characteristics.

The In-Depth Study

The experience of the Department staff in districts offering prekindergarten programs has led to the conclusion that the degree of continuity between prekindergarten, kindergarten, and grades 1 through 3 is generally low. It is difficult to assess the potential effects of a high degree of continuity because of this negative skewing of the degree of continuity. There are many districts with a low degree of continuity but relatively few, if any, with a high degree.

The work of the field coordinator is designed, in part, to help a few districts increase the degree of continuity which presently exists from prekindergarten through grade 3. Thus, a more nearly normal distribution of the degree of continuity will be available, making possible more meaningful studies of the impact of continuity.

Statistical analyses to be conducted to investigate the impact of continuity and related factors will include districts from outside the so-called "in-depth districts" as well as districts from within the group. Therefore, the unrepresentativeness of the "in-depth districts" will not affect the validity of the results but rather will enhance the usefulness

of the results.

The in-depth study will focus on the students in seven districts operating the Experimental Prekindergarten Program. Programs selected include several significant program variations (differences in instructional time and degree of parental involvement); contain sufficient numbers of students to provide an adequate base for the study; exhibit a commitment to follow-through activities in grades K through 3; show promise of maintaining programs through the five-year period of the study; have administrative support and willingness to participate; exhibit a low attrition rate; and are in districts which have reasonably small class size in kindergarten and primary years and are in close proximity to and/or have a good relationship with a Resource Center. An estimated 800 pupils will be included in the in-depth study. Following the same sample over five years will result in the potential use of student profiles and the detection of which programs maintain prekindergarten gains during the following years.

The in-depth study will follow a testing schedule similar to that used in the general study: testing at the beginning and end of the prekindergarten and kindergarten years and at the end of each year through Grade 3.

The in-depth study, compared to the general study, has the following additional components:

1. All 4-year-old students, rather than a sample, will be studied over the five-year period.
2. Observational data will be collected, a process too expensive

and time-consuming for a large-scale study such as the general study. Data will be obtained on the use of language in the classroom, teacher-pupil interactions, and other aspects of classroom activity. Methods of collecting observational data have been developed by Virginia Shipman at Educational Testing Service and such methods will be used as a basis for the prekindergarten observations.

3. Particular attention will be given to articulation (or lack thereof) between prekindergarten and the regular school program in succeeding years.

These factors should make it possible to assess the effectiveness of much more specific variations in the educational experiences of the children than is possible in the general study. A major emphasis will be an attempt to match educational experiences to the characteristics of children. This will allow conclusions beyond answering the question, "Is prekindergarten effective?" It will help answer the question, "Which programs and practices are most effective with children possessing different characteristics?" For planning educational programs, this is a key question.

## DATA TO BE COLLECTED

Data will be collected to describe students' backgrounds, students' development, and the programs to which students are exposed. The data collection system will provide for anonymity of individuals on data transmitted from

the school districts.

Data in several categories are described below.

Data on Student Background

Data collected from all students will include such items as parents' occupation or level of education, whether on public assistance, students' sex, and bilingualness (see Initial Pupil Interview form attached). These data will be obtained from students' records where it is available and supplemented by use of questionnaires or interviews. Such data will be important in determining which program variations are most effective with students having different background characteristics.

Data on Cognitive and Affective Development

Data on student development will be collected in the two studies by a variety of measures. A review of available instruments and their evaluations for early childhood was conducted by Department staff with the following as criteria:

a. a test designed for prekindergarten children
b. standardization, reliability, and validity
c. ease of administration
d. time of administration
e. low language component
f. cost and availability

Each of the measures selected meet these criteria.

A discussion of the rationale for selection, underlying constructs, and child behavior elicited for each data collection device follows.

Cognitive Measures. Cognitive measures are those which assess the intellectual development of the child. The three tests selected are as follows.

## Walker Readiness Test

To control for individual differences, there is a need to measure children's general reasoning at the beginning of the program and again at the beginning of kindergarten when the control group was selected. During the prekindergarten year, the general reasoning measure will be used to analyze the interaction of treatment and general reasoning. At the prekindergarten and kindergarten level, it will be used as a covariate for control and experimental groups.

The Walker Readiness Test (WRT) is a 50-item, non-verbal measure of general reasoning. Each item consists of a set of pictures and the child is requested to choose one of the pictures according to the instructions for that section of the test. The four subtests are:

1. Similarities - From the four option pictures, the child must choose the picture which is identical to the stimulus picture.
2. Differences - From four pictures the child must choose the one that is different with respect to size, characteristic, or category.
3. Numerical - From four collections of objects, the child must select the collection which has the same number of objects as the stimulus collection.
4. Missing Parts - From the four option pictures, the child must choose the one that completes the stimulus picture.

The Walker Readiness Test was designed specifically for low socioeconomic children and was normed on a sample of approximately 12,000 Head Start children. Like the Columbia Test of Mental Maturity (CTMM), it is a nonverbal test in the sense that it presents pictorial stimuli and the child makes a pointing response. Equivalence reliability was calculated at .693 for the urban sample and at .779 for the rural sample. Split-half reliability for both samples was in the high .80's. Construct validity can be inferred from the correlations with other test--Form A of the Walker correlates .791 with the Metropolitan Test and Form B .757 with the same test. Correlations are .807 (Form A) and .708 (Form B) with the Cooperative Preschool Inventory and .677(Form A) and .697 (Form B) with the CTMM. An additional source of validity is the .703 correlation between teacher rankings of a child's readiness for kindergarten and scores on the Walker (Walker, 1969).

Since the Walker has two forms, one source of measurement error, that of using the same form of a test for multiple administration, will be reduced. The test is relatively easy to administer, would require a minimum amount of examiner training, and takes only about 15 minutes testing time per child.

Other Measures Considered But Not Selected

An individualized intelligence test, such as the Stanford-Binet, Wechsler Preschool and Primary Scale of Intelligence, or the McCarthy Scale of Children's Abilities, was rejected because of the cost of administration, the excessive administration time, and the logistics of setting up regional centers for training examiners.

One other general reasoning test, Columbia Test of Mental Maturity, was considered. This test was rejected because all questions were based on the same underlying construct--the child's understanding of differences. If a child answered incorrectly, there was no way of knowing whether this was due to his lack of perceiving important similarities in the stimuli that went together or to his lack of understanding of difference.

## Cooperative Preschool Inventory

The second cognitive measure selected, the Cooperative Preschool Inventory (CPI), was chosen because it was developed to measure achievement in areas regarded as necessary for success in school. The test was originally constructed to evaluate the first Head Start programs. Standardization of the revised edition took place in 11 Head Start centers. Kuder-Richardson Formula 20 reliability coefficients for the age groups range from .86 to .92. For the total standardization sample, the mean number of correct responses increases about 4 raw score points from one age group to the next. The standard error of measurement for the age groups varies from 3.1 to 3.9 raw-score points (Educational Testing Service, 1970). Criterion-based validity is based on correlations of .57 and .60 of scores on the CPI in prekindergarten with third-grade mathematics and reading scores.*

The test contains 64 questions most of which deal with specific behaviors falling into 4 categories: personal-social responsiveness, associative vocabulary, concept activation-numeral, concept activation-sensory. The personal-social responsiveness category measures the child's knowledge about his own personal world and his ability to get along with and respond to communications of another person. (Examples are: What is your first name,

* Personal Communication with Virginia Shipman at ETS.

How old are you?, What do you call [knee] ____?, Put three cans by the box, Put two cans behind the middle box.) The associative vocabulary questions, which are much more like the information and comprehension subtests of the Wechsler Intelligence Scale for Children measures the child's understanding of the connotation of a word. The child demonstrates his understanding by carrying out some action or by associating certain intrinsic qualities of the underlying verbal concept.(Examples are: Who do you go to if you are sick?, When do you eat breakfast?, What does a dentist do?) The concept activation-numerical category measures the child's ability to label quantities in order to make judgments of more or less, and to recognize seriated positions. (Examples are: How many eyes do you have?, How many wheels does a car have?, In a series of chickens, point to the second chicken.) Questions in the concept activation-sensory category measure the child's awareness of certain sensory attributes (shape, size, motion, color) and his ability to execute certain visual-motor configurations. (Examples are: Copy a circle, What color is the crayon [black]?)

The Cooperative Preschool Inventory can be administered by the classroom teacher with a minimum amount of training and requires approximately 15 minutes per child to administer.

## Other Measures Considered But Not Selected

One other test, the Tests of Basic Experiences (TOBE), was considered. The TOBE was rejected because the Inventory samples many more behaviors than any of the specific subtests of the TOBE and requires less time to administer.

## Peabody Picture Vocabulary Test

The third instrument chosen, the Peabody Picture Vocabulary Test (PPVT), was selected because neither the Walker Readiness Test nor the Cooperative Preschool Inventory contains a specific language factor. The PPVT is a measure of receptive language. Each child is presented with a set of pictures and then hears a stimulus word and is instructed to select the picture which is the equivalent of the stimulus word.

The test has two forms which can be used for pre- and posttest measurement and the coefficient of equivalence for 4 year olds is .77 (Dunn, 1965). Validity for the PPVT was demonstrated by the correlation scores on the Stanford Binet (DiLorenzo and Brady, 1968). Correlations between the PPVT and reading in first grade are approximately .40 (Klaus and Starke, 1964). The PPVT can be administered by the classroom teacher and requires approximately 15 minutes per child.

## Other Measures Considered But Not Selected

The Revision of the Smith Vocabulary Test by Williams and McFarland was rejected because of the low ceiling age, 6 years 6 months. Several other verbal language measures reviewed were rejected because they were to begin after pre-school or had a ceiling of kindergarten, both of which would limit their usefulness.

## Summary

The cognitive measures can all be administered by the teacher in relatively little time and are largely nonverbal. All have been normed on prekindergarteners, two of them on disadvantaged groups. Total administration time for the three measures will be approximately 45 minutes.

The reliability and validity data are acceptable for prekindergarten tests.

Affective Measures. In selecting affective measurements, several problems had to be resolved. First, the affective domain is quite broad. In order to demonstrate treatment effects, the instrument chosen had to be sensitive to the proposed goals of the programs. The second problem was whether to use 1) a teacher rating of affective behavior or 2) a projective technique whereby the child responds to questions from an examiner and inferences are made directly or indirectly about the child's self-concept or achievement motivation. Because of the rather low reliability and questionnable validity of the projective technique, the rating scale method was chosen.

Two short rating scales were selected. Both of the instruments, California Preschool Social Competency Scale (CPSCS), and the Classroom Behavior Inventory (CBI), have been favorably reviewed in the literature (Proger, 1974 and Walker, 1974, respectively). Both measures were also used to assess social motives in the ETS Head Start Longitudiual Study.

## Classroom Behavior Inventory

Construct validity of the Classroom Behavior Inventory was investigated using factor analysis which yielded three major dimensions (Schaefer and Aaronson, 1965). The first dimension has been called task-oriented behavior and consists of questions dealing with perseverance and attentiveness. The opposite pole of the dimension typically shows loadings on traits such as distractibility and hyperactivity. A second

factor, labelled hostility, has high loadings on questions concerning resentfulness and quarrelsomeness. The opposite pole of this dimension has highest loadings for considerateness. A third factor, which has been labelled extroversion vs. introversion, measures the child's social initiative.

The instrument consists of 15 scales, is completed by the teacher, and takes approximately 5-10 minutes per child. Ratings are cumulative assessments--the teacher observing the child over a period of time but completing the rating sheet at one sitting. One benefit of using such an instrument is that it will provide the teacher with a frame of reference for making more detailed observations.

California Preschool Social Competency Scale

The California Preschool Social Competency Scale was selected because 1) the four factors it assesses (work habits, personality, communication, and interpersonal relations) complement the Classroom Behavior Inventory and 2) the measure will add several important dimensions.

Parental Attitude Measures. Current research (Bronfenbrenner, 1974) supports the hypothesis that children make greater gains in development when their parents are involved in their educational programs. While each program will log parental involvement both in terms of contact hours and type of activities, another aspect of the evaluation that needs quantifying is the effect of parental involvement on the attitudes of parents towards education and child rearing.

A search of the literature failed to produce a single instrument which would assess both of the following goals of the Experimental Prekindergarten Program in relation to parental involvement, i.e. to improve parental attitude toward education and toward child rearing practices. However, several scales were identified which purport to assess the former goal and two for the latter goal. From those scales identified, the following are the most appropriate:

1. Parental Attitude Scale - assesses the attitude of parent in the treatment of children
2. Educational Attitude Survey - assesses parental attitude toward school

Observational Measures. Observational data will be collected in the in-depth study and will describe education activities of students and their parents. State Education Department staff will obtain or develop instruments for collecting such data.

Documentation. Documentation as a method of evaluation is particularly effective when there is a desire to reveal components of an integrative, developmental program designed to meet inclusive needs of a child and his family. Researchers who are concerned with the continuation of developmental programs have given considerable study to the usefulness of documentation.

Documentation is important to the complete evaluation of the prekindergarten program. Collecting information on each child will

give a more comprehensive picture of each child's strengths. While the product (descriptive data) will give a more complete picture of children and the program, the process will serve as a method for program improvement.

The evaluation team will assist all districts in data collection for the evaluation by providing forms to insure uniformity. Forms which will be forwarded to each center include: background information, parental involvement, social services information, program description, health and nutrition, and attendance. These records will be common to all 53 of the prekindergarten centers. The material will be filed with individual teachers and will be accessible to the evaluation team.

The seven districts chosen for in-depth study will develop records on individual children in the program and use them to evaluate and improve experiences for each child. Selected case studies will be written from these collections of data. In addition to the information collected by standard form, each center will be asked to file a Draw-A-Person test for each child three times each school year. For evaluation purposes the Koppitz (Koppitz, 1968) adaptation of the Goodenough-Harris scoring system of developmental level will be used. The Draw-A-Person sample along with recording of the child's height and weight and chronological age serves to document the child's development in an easy to collect, easy to use record for parents as well as for teachers and evaluators. In the in-depth study schools, the following additional records will be kept to enrich the case study: samples of work (paintings, drawings, dictation or writing), parent conference reports, sociometric data (sociograms and

observations), teacher's record of children's experiences, and record of other assessments.

## DATA ANALYSIS

A variety of statistical tools are available for analyzing the data. Two general statistical tools are described below.

### Statistical Tools

Analysis of Variance and Covariance. Analysis of variance and analysis of covariance are techniques for determining the significance of differences between groups of students on some criterion of performance. These procedures are especially useful in studying the effects of different treatments.

Multiple Regression Analysis. Multiple regression analysis is useful in trying to ascertain the effects of a number of variables on a criterion of performance. The technique will be especially appropriate for studying the effects of variations in programs.

These techniques and associated procedures will be the principal analytical tools used in the evaluation study.

### Design Factors

In summary, the design will have the following factors:

A. Covariates

1. SES variables - SES data together with other data obtained on each child's data sheet
2. Walker Readiness Test - as a measure of general reasoning

B. Treatment Variables

1. Exposure time - Number of contact hours the program provides for the child. Programs differ operationally according to length of day (half day or whole day) and number of days per week (one to five) or any combination of length of day and number of days per week.
2. Program type - Home-based or Center-based
3. Treatment - Prekindergarten or no prekindergarten experience
4. Kindergarten type - Different types of classrooms prekindergarten children are in beginning with kindergarten
5. Teacher Differences - amount and type of teaching experience
6. Amount of parental involvement

C. Criterion Variables

1. Cooperative Preschool Inventory
2. Peabody Picture Vocabulary Test
3. Classroom Behavior Inventory
4. California Preschool Social Competency Scale
5. Parental Attitude Scale
6. Educational Attitude Scale
7. Observational Data (in-depth study only)
8. Draw-A-Person (in-depth study only)

Overall Design

In a general format, the results from the research design can be shown as figure 1. Figure 1 shows no difference between the prekindergarten and

non-prekindergarten groups at the beginning of the prekindergarten year. Measures of cognitive and affective growth are taken at many points in time (seven) through grade 3 on the same children.

Questions 1 and 2 in the plan concern the effect of participating in a prekindergarten program. The gain obtained by participation is obtained in this design by the difference between the planes at point K. Figure 2 manifests this difference in the shaded portion of the figure. This is the effect of the prekindergarten programs on children's development at the kindergarten level (Question 2). The effect at the end of prekindergarten year (Question 1) would be the same space moved downward slightly.

Questions 4, 5, and 6 in the plan concern the effect of prekindergarten over the entire range through the third grade. This influence is manifested by the vertical thickness of the block shown in figure 3. For the purpose of clarity, only the thickness at the back of the block is shaded. The reader can project this thickness to the front of the block. Question 3 asks whether the block has any thickness at all. Questions 4 and 5 additionally ask whether the thickness changes toward the back of the block.

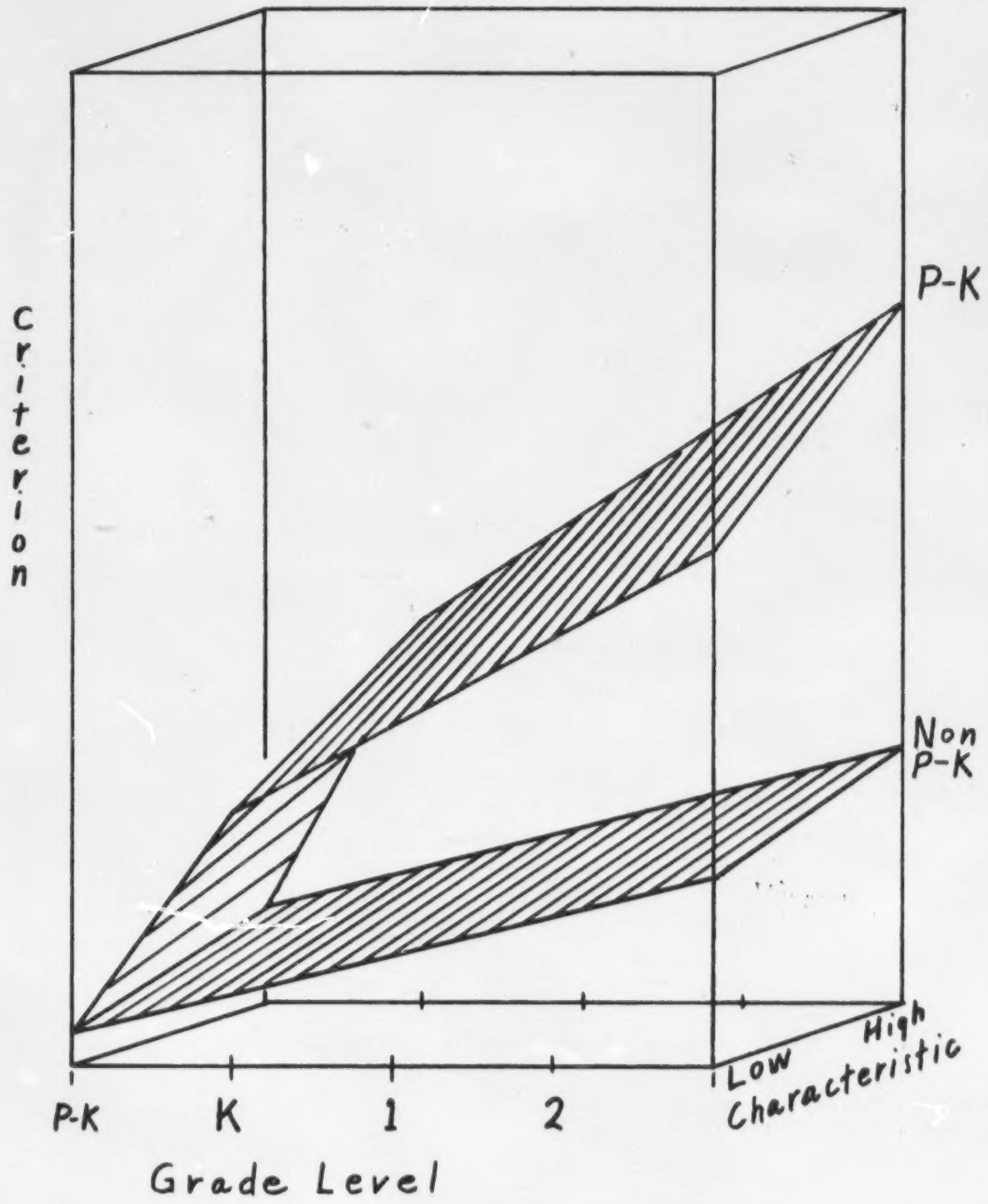


Figure 1. Representation of Results

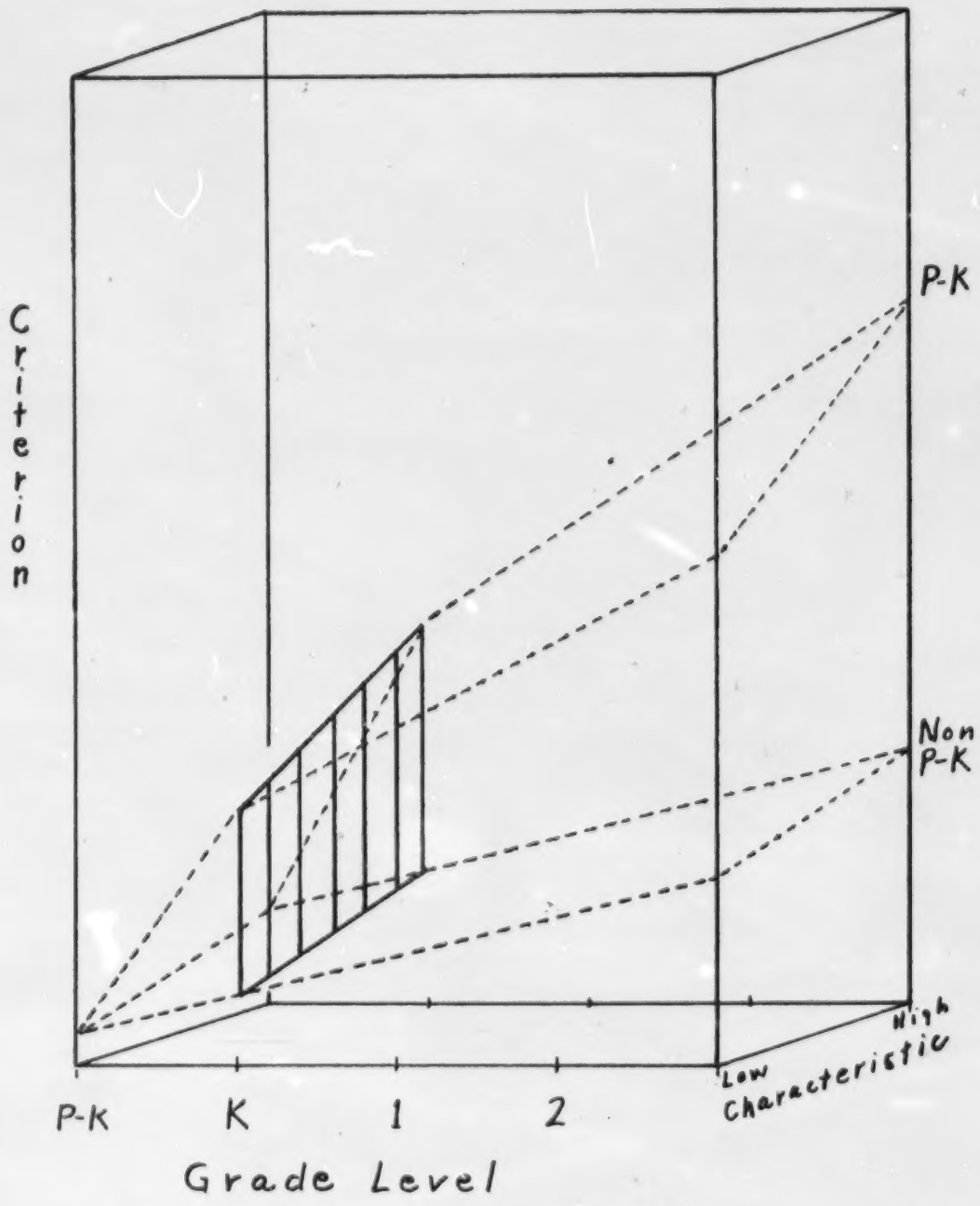


Figure 2. Effect of the Prekindergarten Program At the Kindergarten Level

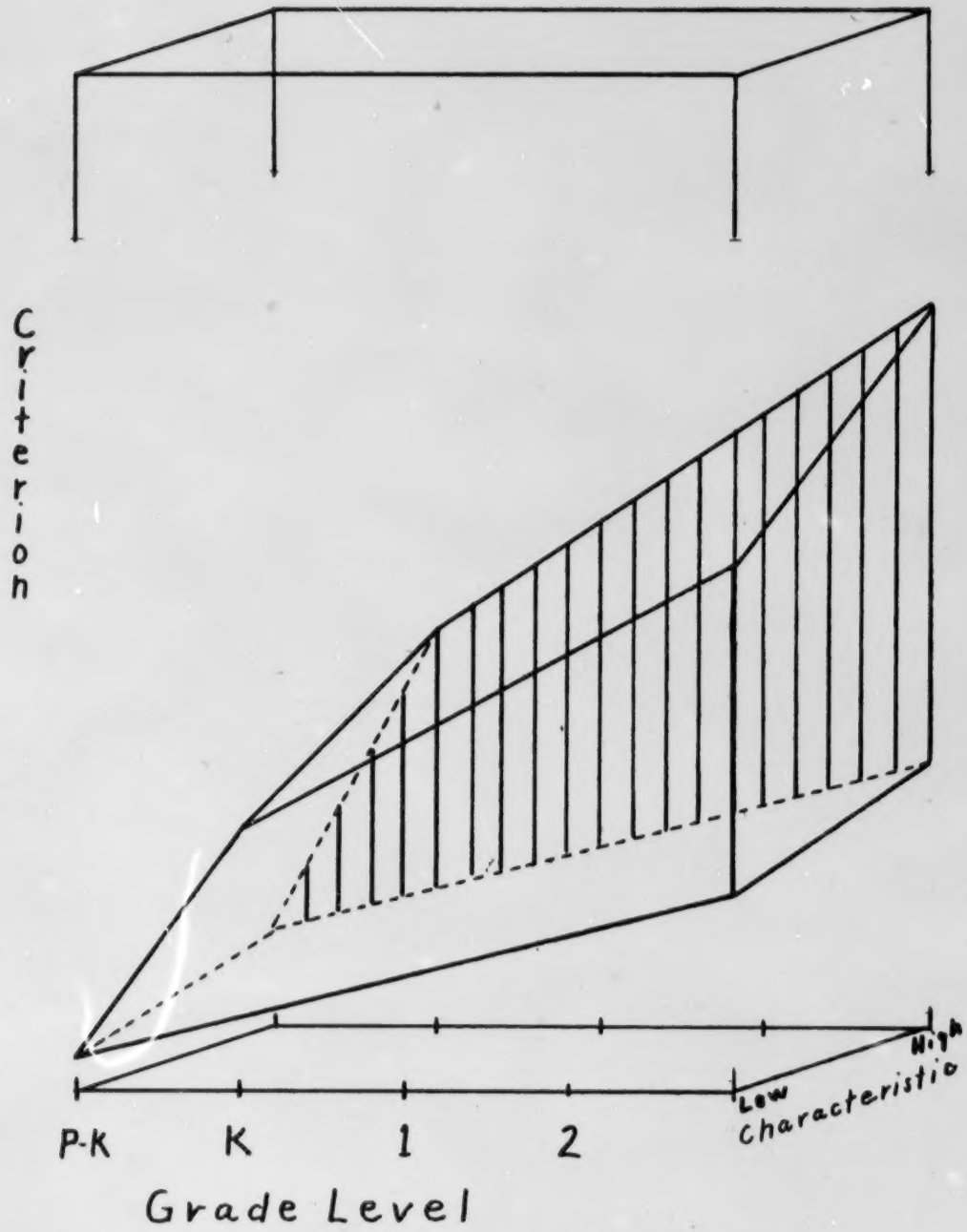


Figure 3. Overall Effect of Prekindergarten Program Manifested By Block Thickness

TIME FRAME

| Fiscal year* | | Activities |
|---|---|---|
| 1975-76: | Summer | Staffing; developing a detailed plan; coordinating with districts; selecting instruments; developing data collection procedures; orienting temporary personnel training test administrators |
| | Fall and winter | Administering fall tests to Wave I** Prekindergarten students; analyzing pretest data |
| 1976-77: | Spring and summer | Administering spring tests to Wave I Prekindergarten students; compiling and analyzing Prekindergarten data; writing first-year report |
| | Fall and winter | Administering fall Kindergarten tests to Wave I and the fall prekindergarten tests to Wave II; compiling data |
| 1977-78: | Spring and summer | Administering spring tests to Wave I Kindergarten students; compiling and analyzing Kindergarten data; administering prekindergarten Wave II spring tests; compiling and analyzing Wave II |
| | Fall and winter | Analyzing Prekindergarten Wave I and II and Kindergarten (Wave I) data; writing interim report; administering prekindergarten Wave III tests |

| Fiscal year* | | Activities |
|---|---|---|
| 1978-79: | Spring and summer | Administering spring tests to Grade 1 students (Wave I); compiling and analyzing Grade 1 data; administering spring tests to prekindergarten Wave III; compiling and analyzing this data |
| | Fall and winter | Analyzing Prekindergarten, Kindergarten, and Grade 1 (Wave I) data; writing interim report; administering tests and analyzing prekindergarten data Wave IV |
| 1979-80: | Spring and summer | Administering spring tests to Grade 2 (Wave I) students; compiling and analyzing Grade 2 data; administering and analyzing test data for prekindergarten Wave IV |
| | Fall and winter | Analyzing Prekindergarten, Kindergarten, Grade 1 and Grade 2 (Wave I) data; writing interim report; administering and analyzing test data for prekindergarten Wave V |
| 1980-81: | Spring and summer | Administering spring tests to Grade 3 (Wave I) students; compiling Grade 3 data; administering and analyzing test for prekindergarten Wave V |
| | Fall and winter | Analyzing five-year data; writing final report |

---

* April 1 to March 31

** Wave I- prekindergarteners in 1975-76 academic year; all testing will include appropriate control groups

Wave II- prekindergarteners in 1976-77 academic year

Wave III- prekindergarteners in 1977-78 academic year

Wave IV- prekindergarteners in 1978-79 academic year

Wave V- prekindergarteners in 1979-80 academic year

BIBLIOGRAPHY

Bronfenbrenner, U. A Report on Longitudinal Evaluations of Preschool Programs. Volume II: Is Early Intervention Effective? Washington, D. C.: Department of Health, Education, and Welfare, Document Number 74-25, 1974.

DiLorenzo, L. and J. Brady, "Use of the Peabody Picture Vocabulary Test With Preschool Children," Psychological Reports, 1968, 22, 247-251.

Dunn, Lloyd M. Peabody Picture Vocabulary Test: Manual. Circle Pines, Minn.: American Guidance Service, Inc., 1965.

Educational Testing Service. Preschool Inventory, Revised Edition - 1970: Handbook. Princeton, N.J.: Educational Testing Service, 1970.

Klaus, R.A. and C. Starke. Experimental Revision of the Peabody Picture Vocabulary Test As A Predictor of First Grade Reading Ability. Unpublished paper. Nashville, Tennessee.: Peabody College, Psychology Department, 1964.

Koppitz, E.M. Psychological Evaluation of Children's Human Figure Drawings. New York: Grune and Stratton, 1968.

Proger, B., "Test Review, No. 17, California Preschool Social Competency Scale," Journal of Special Education, 1974, 8, 391-395.

Walker, D., ed. Socioemotional Measures for Preschool And Kindergarten Children. San Francisco: Josey-Bass, 1973.

Walker, W. "Development of a Readiness Test For Disadvantaged Pre-School Children In The United States. Final Report." Northwest Missouri State College, Maryville, Missouri, 1969.